

University of Groningen

## An analysis of critical-link semantics with variable degrees of justification

Wei, Bin; Prakken, Henry

*Published in:*  
Argument & Computation

*DOI:*  
[10.3233/AAC-160003](https://doi.org/10.3233/AAC-160003)

**IMPORTANT NOTE:** You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

*Document Version*  
Publisher's PDF, also known as Version of record

*Publication date:*  
2016

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Wei, B., & Prakken, H. (2016). An analysis of critical-link semantics with variable degrees of justification. *Argument & Computation*, 7(1), 35-53. <https://doi.org/10.3233/AAC-160003>

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# An analysis of critical-link semantics with variable degrees of justification

Bin Wei <sup>a,\*</sup> and Henry Prakken <sup>b,c</sup>

<sup>a</sup> *School of Administrative Law, Southwest University of Political Science and Law, Chongqing, China*

<sup>b</sup> *Department of Information and Computing Sciences, Utrecht University, Utrecht, The Netherlands*

<sup>c</sup> *Faculty of Law, University of Groningen, Groningen, The Netherlands*

**Abstract.** This paper is to critically examines Pollock’s critical-link semantics with variable degrees of justification. Some possibly counterintuitive consequences of Pollock’s definition of degrees of justification are identified and a modified definition is proposed which avoids these consequences. Then the new solution is applied to the case of so-called presumptive defeat. A second contribution of the paper is to show how the modified semantics can be applied to the *ASPIC*<sup>+</sup> framework: first the *ASPIC*<sup>+</sup> framework is modified to allow for variable degrees of justification and then the modified way to compute these degrees is applied in a new notion of an argument graph.

**Keywords:** John L. Pollock, critical-link Semantics, variable degrees of justification, defeasible reasoning, *ASPIC*<sup>+</sup>

## 1. Introduction

John L. Pollock [11,15] initiated the formal study of argumentation in AI based on his contributions to epistemology, and his approach is still very influential. Pollock introduced the distinction between deductive and defeasible reasons and the corresponding distinction between rebutting and undercutting defeat. Many of today’s approaches are still based on these ideas, such as *ASPIC*<sup>+</sup> [16] and Defeasible Logic Programming [4]. Probably the best-known version of his work is his 1995 book [15]. The semantics of this version was by Jakobovits proved to be equivalent to Dung’s [2] preferred semantics [7]. Later Pollock revised his semantics, for two main reasons: he wanted to enforce the same treatment of odd and even defeat loops (like also Baroni and Giacomin [1]) and he wanted to make the justification status of conclusions a matter of degree. To achieve both aims, Pollock developed his so-called critical-link semantics for his system in [12,13].

Pollock introduced variable justification degrees to account for the so-called “diminishing” effect of attempted defeaters that are weaker than their target. In such cases Pollock wanted to model that the attempted defeaters can still weaken the degree of justification of their target. More generally, the idea that justification of conclusions is not an all-or-nothing affair but a matter of degree is very natural and therefore deserves to be studied. The present paper aims to contribute to such a study by critically examining Pollock’s proposal.

In particular, we will argue that Pollock’s approach in some cases arguably gives counterintuitive outcomes, and we will modify his account in a way that avoids these outcomes. Among other things we

---

\*Corresponding author. E-mail: [srsysj@126.com](mailto:srsysj@126.com).

will apply our new proposal to the issue of “presumptive defeat”, also known as the issue of ambiguity blocking versus ambiguity propagating. Since many of Pollock’s ideas are still used in other frameworks, our contributions will also be relevant for these frameworks. To illustrate this point, we will at the end discuss how Pollock’s ideas and our modifications can be incorporated in the *ASPIC*<sup>+</sup> framework for structured argumentation recently proposed by [9,10,16–18].

This paper is organized as follows. In Section 2, we first summarize Pollock’s 1995 system [15], then discuss why Pollock wanted to change this and then give his critical-link semantics and his definition of gradual justification. In Section 3, we then discuss some arguably counterintuitive outcomes, present our revised definitions and show that they avoid these outcomes. In Section 4, we discuss how to transfer the revised semantic into *ASPIC*<sup>+</sup> framework with gradual degrees. Finally, we conclude in Section 5.

## 2. Semantics

In this section we present Pollock’s critical-link semantics with variable degrees of justification, preceded by a brief overview of his multiple-assignment semantics.

### 2.1. Basic features

There are several constant features in Pollock’s work on defeasible reasoning. Reasoning proceeds from a knowledge base of classical-logic formulas by chaining reasons into inference graphs, where all reasons are either deductive or defeasible. Only applications of defeasible reasons can be defeated, and there are two kinds of defeaters: *rebutting* defeaters attack the conclusion of a defeasible inference by favoring a conflicting conclusion, while *undercutting* defeaters attack the defeasible inference itself, without favouring a conflicting conclusion.

More precisely, Pollock assumed as given a knowledge base of first-order formulas and two sets of deductive and defeasible reasons, which technically are inference rules. Pollock then considered arguments, which are sequences of *argument lines* (in later work, he speaks of sequences of nodes in an inference graph). Each argument line contains a proposition, the reason applied to infer the proposition (where this reason can also be that it is taken from the knowledge base), the set of preceding lines from which the proposition is inferred, and the line’s strength. Both elements from the knowledge base and reasons have a numerical strength, which are used to compute the strength of an argument in a way further explained below. The strength of an element  $\varphi$  of the knowledge base is below written as  $\delta(\varphi)$  while the strength of a reason  $r$  will be written as  $\rho(r)$ .

**Definition 1.** An argument line is a tuple  $(\varphi, r, L, s)$ , where  $\varphi$  is a proposition,  $r$  is the reason applied to infer  $\varphi$ ,  $L$  is the set of preceding lines from which  $\varphi$  is inferred, and  $s$  is the line’s strength.

Below the strength of argument line  $l$  will sometimes be written as  $s(l)$ .

**Definition 2.** An argument line  $(\varphi, r, L, s)$  defeats an argument line  $(\varphi', r', L', s')$  iff  $r'$  is a defeasible rule, and  $s \geq s'$ , and  $\varphi = \neg\varphi'$  or  $\varphi = \neg r'$  (here  $\neg r$  is shorthand for saying that the antecedents of rule  $r$  do not support its consequent). Moreover, an argument  $A$  defeats an argument  $B$  iff a line of  $A$  defeats a line of  $B$ .

In his account of the strength of arguments, Pollock rejected the probability calculus and initially proposed a *weakest link principle* as the underlying principle for computing the strength of an argument. He defined the argument strength of a defeasible argument as the minimum of the strengths of the defeasible reasons employed in it and the degrees of justification of its premises.

**Definition 3.** For any argument line  $l = (\varphi, r, L, s)$  (where  $L = \{l_1, \dots, l_n\}$ ) the strength  $s(l)$  is inductively defined as follows:

- If  $l$  takes  $\varphi$  from the knowledge base, then  $s(l) = \delta(\varphi)$ .
- Otherwise,  $s(l) = \min\{\rho(r), s(l_1), \dots, s(l_n)\}$ .

With respect to accrual of arguments for the same conclusion, Pollock proposed that if there are two or more separate undefeated arguments for a conclusion, the degree of justification for the conclusion is the maximum of the strengths of these arguments.

## 2.2. Multiple assignment semantic

In [15] Pollock considered inference graphs, where the nodes represent the conclusions of argument lines, support-links tie nodes to the nodes from which they are inferred from  $L$  or the conclusions of the argument lines, and defeat-links indicate defeat relations between nodes. These links relate their roots to their targets. The root of a defeat-link is a single node, while the root of a support-link is a set of nodes. He then proposed a labeling approach to define the justification status of nodes and propositions.

**Definition 4.** A node of the inference-graph is *initial* iff its node-basis and list of node-defeaters is empty, where

- The *node-basis* of a node is the set of roots of its support links.
- The *node-defeaters* are the roots of the defeat links having the node as their target.

**Definition 5.** An assignment  $\sigma$  of defeated and undefeated to a subset of the nodes of an inference-graph is a *partial status assignment* iff:

- $\sigma$  assigns *undefeated* to any initial node;
- $\sigma$  assigns *undefeated* to a non-initial node  $\alpha$  iff  $\sigma$  assigns *undefeated* to all the members of the node-basis of  $\alpha$  and  $\sigma$  assigns *defeated* to all node-defeaters of  $\alpha$ ;
- $\sigma$  assigns *defeated* to a non-initial node  $\alpha$  iff either  $\sigma$  assigns *defeated* to a member of the node-basis of  $\alpha$  or  $\sigma$  assigns *undefeated* to a node-defeater of  $\alpha$ .

**Definition 6.** Assignment  $\sigma$  is a *status assignment* iff  $\sigma$  is a partial status assignment and  $\sigma$  is not properly contained in any other partial status assignment.

Finally, Pollock defines the justification status of an argument as follows:

**Definition 7.** A node  $\alpha$  of an inference graph is *undefeated* iff every status assignment to the inference graph assigns *undefeated* to  $\alpha$ ; otherwise  $\alpha$  is *defeated*.

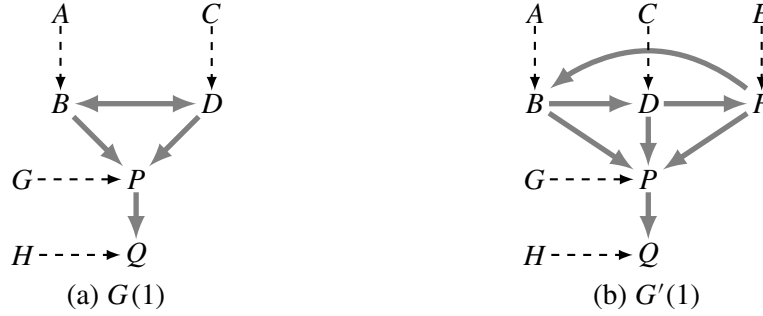


Fig. 1. Inference graphs.

Although the multiple assignment semantics produces the intuitively correct answer for many complicated inference-graphs, Pollock was still not satisfied with it, since it can produce different outcomes for inference-graphs containing odd-length and even-length defeat cycles. See, for example, the graphs in Fig. 1 (dotted arrows indicate support links and full arrows indicate defeat links):

In  $G(1)$ , there are two status assignments: in the first assignment  $B$  is assigned defeated while  $D$  is assigned undefeated while in the second assignment  $B$  is assigned undefeated while  $D$  is assigned defeated. In both assignments  $P$  is defeated by an undefeated defeater, so in both assignments  $P$  is defeated and hence  $Q$  is undefeated. Hence according to Definition 7,  $P$  and  $Q$  respectively have the justification status defeated and undefeated. By contrast, in  $G'(1)$  there is only one status-assignment, namely an assignment that assigns no status to any node of  $B, D, F, P$ , or  $Q$ . So according to Definition 7, both  $P$  and  $Q$  have the justification status defeated.

Pollock believed that the difference of statuses of  $Q$  in these two graphs is a clear counterexample to all existing semantics. In order to avoid this problem, he proposed a new so-called critical-link semantics.

### 2.3. Critical-link semantics on status of justification

The core idea of critical-link semantics [12,13] for solving the above problem is to build new inference-graphs as subgraphs of the original inference graph and assign various statuses to initial nodes in different cases. This idea is formally defined as follows:

**Definition 8.** An *inference/defeat-path* from a node  $\varphi$  to a node  $\theta$  is a sequence of support-links and defeat-links such that

- $\varphi$  is a root of the first link in the path;
- $\theta$  is the target of the last link in the path;
- the root of each link after the first member of the path is the target of the preceding link;
- the path does not contain an internal loop, i.e., no two links in the path have the same target.

For simplicity, we define the *inference/defeat-path* from a node  $\varphi$  to a node  $\theta$  as a sequence of support-links and defeat-links  $\langle L_1, \dots, L_n \rangle$ , such that (1) node  $\varphi$  is the initial node in path; (2) there exist nodes  $\psi_1, \dots, \psi_{n-1}$ , such that  $L_1 = (\varphi, \psi_1)$ ,  $L_i = (\psi_{i-1}, \psi_i)$ , and  $L_n = (\psi_{n-1}, \theta)$ , where  $(\varphi, \psi)$  means the support-link or defeat link from node  $\varphi$  to node  $\psi$ .

**Definition 9.** A node  $\theta$  of an inference graph is  $\varphi$ -*dependent* iff there is an inference/defeat-path from  $\varphi$  to  $\theta$ .

In  $G(1)$ ,  $B/D$  is  $D/B$ -dependent, while in  $G'(1)$ ,  $B$  is  $D$ -dependent and  $F$ -dependent,  $D$  is  $F$ -dependent and  $B$ -dependent and  $F$  is  $D$ -dependent and  $B$ -dependent.

**Definition 10.** A *circular inference/defeat-path* from a node  $\varphi$  to itself is an inference/defeat-path from  $\varphi$  to  $\varphi$  via a defeater of  $\varphi$ .

In  $G(1)$  there exist two circular defeat paths  $\langle (B, D), (D, B) \rangle$  and  $\langle (D, B), (B, D) \rangle$ , while in  $G'(1)$  there exist three circular defeat paths  $\langle (B, D), (D, F), (F, B) \rangle$ ,  $\langle (D, F), (F, B), (B, D) \rangle$  and  $\langle (F, B), (B, D), (D, F) \rangle$ .

**Definition 11.** A *circular inference/defeat-path* is an *undefeated* circular path iff it is a circular path and there is no defeat link outside the path having a node in the circular path as its target.

In  $G(1)$  and  $G'(1)$ , all of the circular paths are *undefeated* circular path. Consider next three nodes  $A$ ,  $B$  and  $C$ , where  $A$  defeats  $B$ ,  $B$  defeats  $C$  and  $C$  defeats  $B$  (ignore the node-basis). From the above definition it follows that the circular paths  $\langle (B, C), (C, B) \rangle$  and  $\langle (C, B), (B, C) \rangle$  are not *undefeated* circular paths.

**Definition 12.** A defeat-link is  $\varphi$ -critical iff it is a member of some minimal set of defeat-links such that removing all the defeat-links in the set suffices to cut all the circular inference/defeat-paths from  $\varphi$  to  $\varphi$ .

In  $G(1)$ ,  $(B, D)$  and  $(D, B)$  are both  $B$ -critical and  $D$ -critical, since  $\{(B, D)\}$  and  $\{(D, B)\}$  are the minimal sets such that removing  $(B, D)$  and  $(D, B)$  suffices to cut the circular defeat paths  $\langle (B, D), (D, B) \rangle$  and  $\langle (D, B), (B, D) \rangle$ . Likewise, in  $G'(1)$ ,  $(B, D)$ ,  $(D, F)$  and  $(F, B)$  are all both  $B$ -critical,  $D$ -critical and  $F$ -critical, since  $\{(B, D)\}$ ,  $\{(D, F)\}$  and  $\{(F, B)\}$  are the minimal sets such that removing their sole member suffices to cut the circular defeat paths of which they are part.

**Definition 13.** If  $\varphi$  is a node of an inference graph  $G$ , then  $G_\varphi$  is the inference graph that results from

1. deleting all  $\varphi$ -critical defeat-links from  $G$  and
2. making all members of the node-basis of  $\varphi$  initial nodes in  $G_\varphi$  and
3. making all  $\varphi$ -independent nodes initial-nodes in  $G_\varphi$  with stipulated defeat-statuses the same as their defeat-statuses in  $G$ .

Note that Pollock thus in fact modified the default justification status “undefeated” of initial nodes by stipulating that some initial nodes in a newly-constructed inference graph are defeated, instead of being automatically undefeated. For example, graph  $G(2)$  in Fig. 2 displays  $G_B$  (and also  $G_D$ ). In this graph, both  $B$  and  $D$  are also initial nodes but they are now stipulated to be defeated, since they both had the justification status ‘defeated’ in the original graph  $G(1)$ . Likewise, graph  $G'(2)$  in Fig. 2 displays  $G_B$  (and also  $G_D$  and  $G_F$ ). In this graph, all of  $B$ ,  $D$  and  $F$  are initial nodes but they are now stipulated to be defeated, since they all had the justification status ‘defeated’ in the original graph  $G'(1)$ .

Pollock then redefined his computation principles as follows:

**Definition 14** (Justification status for critical-link semantics). The critical-link semantics consists of two rules:

1. Initial nodes are *defeated* when they are newly initial nodes and belong to the node-basis of a node in an undefeated circular path. Otherwise, initial nodes are *undefeated*.

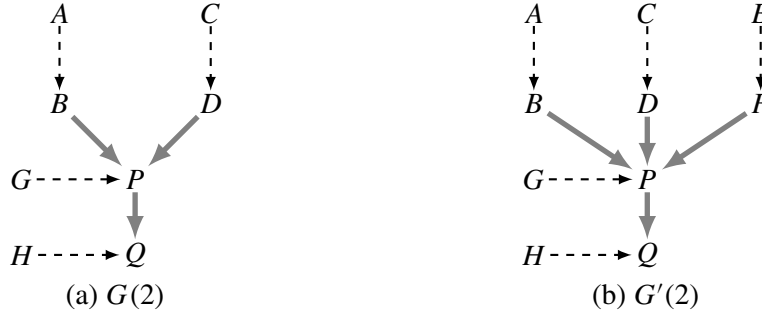


Fig. 2. Renewed inference graphs.

2. A non-initial node  $\varphi$  is *undefeated* in an inference-graph  $G$  iff all members of the node-basis of  $\varphi$  are *undefeated* and any defeater for  $\varphi$  is *defeated* in  $G_\varphi$ .

Note that thus in his critical-link semantics Pollock does not any more consider (possibly multiple) status assignments to a graph but directly defines the justification status of nodes.

In our example this definition yields that node  $Q$  has in both  $G(2)$  and  $G'(2)$  the same status *defeated*. First, all the statuses of nodes  $A, C$  in  $G(2)$  and  $A, C, E$  in  $G'(2)$  are *undefeated*, since they were already initial nodes in the original graphs. Then  $B, D$  in  $G(2)$  and  $B, D, F$  in  $G'(2)$  are *defeated*, since they are newly initial nodes stipulated to be defeated. Consider the statuses of node  $P$  in the new semantics: there are no  $P$ -critical defeat-links in Fig. 2, so removing  $P$ -critical defeat-links leaves the original graph unchanged.  $P$  is then *undefeated*. Then, because there are no  $Q$ -dependent defeat-links in Fig. 2,  $Q$  is *defeated* in both graphs. Note also that this result reflects Pollock' intuition that nodes that are defeated should not have power to defeat other nodes. We will discuss this further below in Section 3.4, when we discuss Pollock' account of so-called "presumptive defeat".

#### 2.4. Critical-link semantics with variable degrees of justification

We next discuss how Pollock uses his critical-link semantics [12] to define variable degrees of justification. A main motivation of the idea that propositions should have variable degrees of justification is Pollock' notion of a *diminisher*. A diminisher is a defeater of a node that is weaker than its target. Note that the notion of defeater here is different from the notion defined in Definition 2, since the new version of the defeat relation does not depend on the strength of arguments anymore. Thus from now on we assume that the condition  $s \geq s'$  in Definition 2 is deleted and we will refer to the defeat notion of Definition 2 as s-defeat (for 'strength defeat'). Now in the above definitions, such a diminisher has no effect on the justification status of its target, since that status is an all-or-nothing affair. However, in Pollock's new approach, the diminisher is still able to diminish the *degree* of justification of its target.

To see how a diminisher works, see again  $G(1)$  in Fig. 1. For simplicity, we assume that the strengths of reasons are at least as great as the degrees of justification of the initial nodes, so that the strength of reasons can be ignored. If the degree justification of node  $C$  is greater than node  $A$ , then the degree justification of node  $D$  is greater than node  $B$ . In multiple assignment semantic,  $B$  is assigned defeated/undefeated and  $D$  is then assigned undefeated/defeated. However, the idea of the new approach is that  $B$  still diminishes  $D$ , that is, the degree of justification of  $D$  is lower than the degree of justification of  $C$ .



For the sake of the mathematics of diminishers, Pollock proposed that there exists a function  $\diamond^1$  such that given two argument lines that rebut one another, if their strengths are  $x$  and  $y$ , the degree of justification for the conclusion of the former is  $x \diamond y$ , while the degree of justification for conclusion of  $y$  is  $y \diamond x$ . He assumed that

the degree of justification can be measured using real numbers, possibly augmented with  $\infty$ , i.e., ‘the extended real numbers’. More precisely, the degrees of justification fall in some interval  $[o, \theta]$ , where  $0 \leq o \leq \theta \leq \infty$ .  $o$  corresponds to no justification, and  $\theta$  to perfect justification, presumably only possible for necessary truths.

Then Pollock defined mathematical properties of  $\diamond$  and its representation as follows:

**Definition 15** (Mathematics of  $\diamond$ ).

- (A1)  $\diamond$  is continuous on the interval  $[o, \theta]$ .
- (A2) If  $\theta > \alpha > \beta > o$ , then  $\alpha > \alpha \diamond \beta > o$ .
- (A3) If  $\theta > \alpha > \beta > \gamma > o$ , then  $\alpha \diamond \beta < \alpha \diamond \gamma$  and  $\alpha \diamond \gamma > \beta \diamond \gamma$ .
- (A4) If  $\theta \geq \alpha \geq \beta \geq o$ , then  $\beta \diamond \alpha = o$ .
- (A5) If  $\theta \geq \alpha > o$ , then  $\alpha \diamond o = \alpha$ .
- (A6) If  $\theta > \alpha$  and  $\beta$  and  $\gamma$  are in  $[o, \theta]$ , then  $(\alpha \diamond \beta) \diamond \gamma = (\alpha \diamond \gamma) \diamond \beta$ .

It should be noted that (A2) means that an argument line with strength  $\beta$  diminishes the other argument line with strength  $\alpha$ , without completely defeating it. To the contrary, (A4) means that an argument line with strength  $\alpha$  completely defeats the other argument line with strength  $\beta$ . We call that the latter argument line in (A2) is incompletely defeated and the latter argument line in (A4) is completely defeated. Moreover, Pollock proved that if (A1)–(A6) hold, then  $\diamond$  has a very simple representation, namely  $\sim$ , which is an instantiation of the operator.

**Definition 16** (Representation of  $\diamond$ ).

$$x \sim y = \begin{cases} x - y & \text{if } y < x < \infty, \\ 0 & \text{otherwise.} \end{cases} \quad (1)$$

Pollock then gave his definition of the computation of degrees of justification.

**Definition 17** (Computation of degree of justification).

- (DJ1) If  $P$  is inferred from the basis  $\{B_1, \dots, B_n\}$  in an inference-graph  $G$  in accordance with a reason of strength  $\rho$ ,  $D_1, \dots, D_k$  are the defeaters for  $P$ , and no  $D_i$  is  $P$ -dependent, then  $J(P, G) = \min\{\rho, J(B_1, G), \dots, J(B_n, G)\} \sim \max\{J(D_1, G), \dots, J(D_k, G)\}$ .
- (DJ2) If  $P$  has  $P$ -dependent defeaters  $D_1, \dots, D_k$  in  $G$ , then  $J(P, G) = J(P, G_P) \sim \max\{J(D_1, G_P), \dots, J(D_n, G_P)\}$ .
- (DJ) If  $P$  is inferred from the basis  $\{B_1, \dots, B_n\}$  in an inference-graph  $G$  in accordance with a reason of strength  $\rho$ ,  $D_1, \dots, D_k$  are the  $P$ -independent defeaters for  $P$ , and  $D_{k+1}, \dots, D_m$  are the  $P$ -dependent defeaters of  $P$ , then  $J(P, G) = \min\{\rho, J(B_1, G), \dots, J(B_n, G)\} \sim [\max\{J(D_1, G), \dots, J(D_k, G)\} + \max\{J(D_{k+1}, G_P), \dots, J(D_m, G_P)\}]$ .

<sup>1</sup>Pollock added the mathematical analysis in his extended version of [12], see <http://johnpollock.us/ftp/PAPERS/Degrees.pdf>.



Principle DJ1 is the computation for the nodes not in a circular path, while DJ2 is for the nodes in a circular path and DJ unites these two principles for “collaborative defeat”, where the nodes are defeated by both node-dependent defeaters and node-independent defeaters. Note that in both cases, if  $P$  is taken from the knowledge base, then the degree of justification of  $P$  equals  $\delta(P)$  (the strength of  $P$ ).

Consider by way of example node  $B$  in  $G(1)$  and let  $\delta(A) = 0.8$  while  $\delta(C) = 0.5$  (recall that by assumption all reasons have strength  $> 0.8$  and can therefore be ignored). Then according to DJ2 we have that  $J(B, G(1)) = J(B, G(2)) \sim J(D, G(2)) = 0.8 \sim 0.5 = 0.8 - 0.5 = 0.3$ .

### 3. Problem cases and modifications

In this section, we discuss some possible problems of Pollock’s critical-link semantics with variable degrees of justification, by analyzing some problem cases.

#### 3.1. Problem case on diminishers

The first problem concerns some arguably counter-intuitive consequences of the mathematical properties and representation of the function  $\diamond$ . We present an example and discuss why the outcomes may be counter-intuitive, and then modify some properties of  $\diamond$  and choose another definition for  $\sim$  to represent  $\diamond$ .

Consider rebutting defeaters in Fig. 3. Let  $P$  be “Jones says that it is not raining”,  $R$  be “Smith says that it is raining”, and  $Q$  be “It is raining”. Let us first assume that Smith and Jones are equally reliable. Then according to Pollock both  $Q$  and  $\neg Q$  should be ‘completely defeated’, since both  $Q$  and  $\neg Q$  should have degree 0, when their strengths are equal. If we apply Definition 17 and again assume that the degrees of justification of the initial nodes are at least as great as the strengths of reasons, then we have  $J(Q, G) = J(\neg Q, G) = 0$ . Assume next that Smith is much more reliable than Jones: then  $Q$  s-defeats  $\neg Q$  while  $\neg Q$  diminishes  $Q$ : by Definition 17 we have  $J(\neg Q, G) = 0$  and  $J(Q, G) = J(R, G) \sim J(P, G) > 0$ .

The arguably counter-intuitive consequence is that node  $\neg Q$  has in both cases the same degree of justification, namely, 0, while yet in the second case the degree of justification of  $Q$  is higher than in the first case. Thus intuitively, in the second case,  $\neg Q$  is s-defeated by  $Q$  with higher strength, so it is intuitive to accept the result that  $\neg Q$  loses the competition with  $Q$  (if we assume that “loses” means that justification equals 0). However, in the first case, if  $\neg Q$  is diminished by  $Q$  with equal strength, then it is counter-intuitive to acknowledge that  $\neg Q$  loses the competition, since its strength is equal to  $Q$ , not lower than  $Q$ . In other words, although node  $\neg Q$  is in the first case not accepted, it is still much more reliable than in the second case. Thus the degrees of justification of nodes in cases of symmetric s-defeat should be greater than the ones in cases of asymmetric s-defeat. Moreover, the first case is similar to

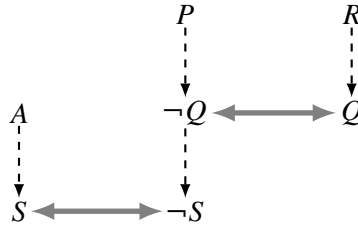


Fig. 3. Inference graph of Presumptive defeat.

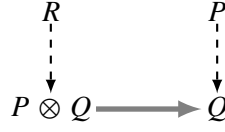


Fig. 4. Inference of Undercutter.

“zombie arguments” [8]: although the arguments are s-defeated, they can still affect other arguments. In other words, the node  $\neg Q$  in the first case still has ability to attack or support other nodes, but the node  $\neg Q$  in the second case does not. So it is necessary to make a difference between the degrees of justification of nodes in these two cases.

### 3.2. Problem case on “presumptive defeat”

The previous point can be further developed in a discussion of ambiguity blocking vs. ambiguity propagating (by Pollock called “presumptive defeat”). Consider again Fig. 3 but let now  $Q$  stand for “Rain was predicted by the morning weather forecast”,  $P$  for “Jones says that no rain was predicted by the morning weather forecast”,  $R$  for “Smith says that rain was predicted by the morning weather forecast”,  $S$  for “It will rain” and  $A$  for “rain was predicted by the afternoon weather forecast”. Suppose again that the reason strengths are at least as great as those of the initial nodes and suppose that  $P$  and  $R$  are equally strong. Then according to Pollock’s new approach the degree of justification of all of  $Q$ ,  $\neg Q$  and  $\neg S$  equals 0, so that  $\neg S$  cannot diminish or “completely defeat”  $S$ . However, according to Section 3.1 the degrees of justification of  $Q$  and  $\neg Q$  should be greater than 0, and this has the consequence that  $\neg S$  potentially has the force to diminish or even completely defeat  $S$ .

### 3.3. Problem case on undercutters

Next we discuss a problem of the computation principle DJ by arguing that it gives an unnatural treatment of the effect of undercutters on the degree of justification of their target. Consider Fig. 4 and let  $P$  be “Jones says that it is raining” and  $Q$  be “It is raining”,  $R$  be “Smith says that Jones always lies” and  $P \otimes Q$  be “Jones is lying” means “ $P$  does not guarantee  $Q$ ”. Note that node  $P \otimes Q$  attacks the connection between node  $P$  and node  $Q$ , so the strength of node  $P \otimes Q$  should arguably directly weaken the strength of the reason from  $P$  to  $Q$  and only indirectly weaken the strength of node  $Q$ . In other words, the strength of an undercutting node should be in comparison with the strength of the reason it undercuts rather than with the strength of the node it attacks. However, in Pollock’s definitions this is not the case.

### 3.4. Modified definition of representation

In his final paper [14], Pollock reconsidered the problem of degrees of justification. He measured degrees of justification using numbers in the interval  $[0, 1]$ , for which reason we henceforth choose the scale as  $[0, 1]$ . From assumptions (A2) and (A4) it is clear that Pollock meant to design the function to distinguish between diminishers that diminish nodes without completely defeating them and diminishers that completely defeat their targets. However, we claim that some assumptions on the mathematical properties of the operator are counter-intuitive and should be revised in order to avoid the above problems.

Firstly, according to the above analysis on diminishers and “presumptive defeat”, assumption (A4) should be modified as follows:

(A4) If  $\theta \geq \alpha > \beta \geq o$ , then  $\beta \diamond \alpha = o$ .

(A4') If  $\theta \geq \alpha = \beta \geq o$ , then  $\beta \diamond \alpha \geq o$ .

These two revised assumptions say that the degrees of justification of nodes in defeat cycles (where defeat is defined as in Definition 2) should be not less than 0 if we choose the scale as  $[0, 1]$ .

Secondly, Pollock wanted that the function for degrees of justification is continuous for both merely diminishing nodes and completely defeating nodes. However, according to the above analysis of diminishers, the degree of justification of a diminished node reduces to real number 0 when the strength of the completing defeating node is approaching to the strength of the diminished node. Then the degree of justification of the diminished node would be definitely greater than 0 in accordance with (A4') if the strengths of the rebutting defeaters are equal. Therefore, the representation is not continuous on the whole interval  $[0, 1]$ , since any point  $(x_0, y_0)$  that satisfies  $x_0 = y_0$  would be a discontinuous point.

Thirdly, the degree of justification for a diminished node should be the strength of this node decremented by an amount determined by the strength of the diminishing node. Moreover, the strength of a node as conclusion is determined by the strength of its reason and the strength of its node as premise. Rebutting defeaters or undercutting defeaters can both act as diminishers but their influences on diminished nodes are different. Undercutting defeaters weaken the strength of the reason they attack, while rebutting defeaters directly weaken the strength of the node as conclusion. Therefore, the order in which undercutting defeaters and rebutting defeaters as diminishers are applied to an argument makes a difference to the degrees of justification, and this in turn means that (A6) is invalid.

In sum, our analysis in Sections 3.1–3.3 makes that assumption (A4) must be modified while assumptions (A1) and (A6) cannot hold. We now define a new representation  $\sim$  for operator  $\diamond$ , which matches the above-revised assumptions. Let us define:

$$x \sim y = \begin{cases} x(1 - y) & \text{if } 0 \leq y \leq x \leq 1, \\ 0 & \text{otherwise.} \end{cases} \quad (2)$$

Next we prove that the new function satisfies the revisions of Pollock's assumptions:

(A1)  $x \sim y$  is not continuous on the interval  $[0, 1]$ .

**Proof.** For any point  $(x_0, y_0)$ , where  $x_0 = y_0 > 0$ . If  $0 \leq y \leq x \leq 1$ , then  $\lim_{(x,y) \rightarrow (x_0,y_0)} x \sim y = x_0(1 - x_0) \geq 0$ . If  $0 \leq x < y \leq 1$ , then  $\lim_{(x,y) \rightarrow (x_0,y_0)} x \sim y = 0$ . Hence,  $x \sim y$  is not continuous on the interval  $[0, 1]$ .  $\square$

(A2) If  $1 > x > y > 0$ , then  $x > x \sim y > 0$ .

**Proof.** If  $1 > x > y > 0$ , then  $x - x(1 - y) = xy$ , hence  $x > x \sim y$ . We further have  $x \sim y = x(1 - y) > 0$ , since  $x > 0$  and  $1 - y > 0$ .  $\square$

(A3) If  $1 > x > y > z > 0$ , then  $x \sim y < x \sim z$  and  $x \sim z > y \sim z$ .

**Proof.** If  $1 > x > y > z > 0$ , then  $x(1 - z) - x(1 - y) = x(y - z) > 0$ , hence  $x \sim y < x \sim z$ , and  $x(1 - z) - y(1 - z) = (x - y)(1 - z) > 0$ , hence  $x \sim z > y \sim z$ .  $\square$

(A4) If  $1 \geq x > y \geq 0$ , then  $y \sim x = 0$ .

It's easy to directly prove this property by definition of function (2).

(A4') If  $1 \geq x = y \geq 0$  then  $x \sim y \geq 0$ .

**Proof.** Actually, we can get a precise result:  $0.25 \geq x \sim y \geq 0$ . If  $1 > x = y \geq 0$ , then  $x \sim y = x(1 - y) = x(1 - x) = -(x - 0.5)^2 + 0.25$ , hence  $0.25 \geq x \sim y \geq 0$ .  $\square$

(A5) If  $1 \geq x > 0$ , then  $x \sim 0 = x$ .

**Proof.** If  $1 \geq x > 0$ , let  $y = 0$ , then  $x \sim y = x(1 - y)$ , hence  $x \sim 0 = x$ .  $\square$

### 3.5. Modified definition of variable degrees of justification

The revised idea for the problem case of undercutters is that the degree of justification of node  $P$  equals the minimum of the strength of reason after being diminished and the degrees of justification of its premises. Then the computation can be modified as follows: (DJ1) If  $P$  has  $P$ -independent defeaters  $D_1, \dots, D_k$  in  $G$  and has no  $P$ -dependent defeaters, then  $J(P, G) = \min\{(\rho \sim \max\{J(D_1, G), \dots, J(D_k, G)\}), J(B_1, G), \dots, J(B_n, G)\}$ .

We next discuss the case where a node  $P$  is defeated by both  $P$ -dependent defeaters and  $P$ -independent defeaters. We propose that these two kinds of defeaters can unite to defeat node  $P$  with a double counting, but computing it with  $P$ -independent defeaters firstly and then continue to compute it with  $P$ -dependent defeaters. The final computation can be modified as follows:

**Definition 18** (Variable degrees of justification). If  $P$  is inferred from the basis  $\{B_1, \dots, B_n\}$  in an inference-graph  $G$  in accordance with a reason of strength  $\rho$ ,  $D_1, \dots, D_k$  are the  $P$ -independent defeaters for  $P$ , and  $D_{k+1}, \dots, D_m$  are the  $P$ -dependent defeaters of  $P$ , then  $J(P, G) = \min\{(\rho \sim \max\{J(D_1, G), \dots, J(D_k, G)\}), J(B_1, G), \dots, J(B_n, G)\} \sim \max\{J(D_{k+1}, G_P), \dots, J(D_m, G_P)\}$ .

For instance, in Fig. 5, node  $\neg S$  is  $S$ -dependent, node  $S$  is  $\neg S$ -dependent and node  $Q \otimes S$  is  $S$ -independent. Let  $J(P, G) = 0.15$ ,  $J(Q, S) = J(R, G) = 0.8$  and the reasons are equally strong:  $\rho = 0.9$ . Then  $J(Q \otimes S, G) = 0.8$ ,  $\rho \sim J(Q \otimes S, G) = 0.18$ ,  $J(\neg S, G_S) = 0.15$ ,  $J(S, G) = \min\{\rho \sim J(Q \otimes S, G), J(Q, G)\} \sim J(\neg S, G_S) = 0.18 \sim 0.15 = 0.153$ , and  $J(S, G_S) = 0.18$ ,  $J(\neg S, G) = \min\{\rho, J(P, G)\} \sim J(S, G_{\neg S}) = 0.15 \sim 0.18 = 0$ .

### 3.6. Solution to the problem cases

We now show that the new definition avoids the arguably counterintuitive outcomes we described above. We do this by analyzing the example of presumptive defeat, which includes the problem case

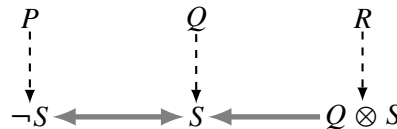


Fig. 5. Inference graph of Collective defeaters.

of diminishers. Consider again the example in Fig. 3. In the multiple-assignment semantics of Definition 5,  $\neg Q$  has the ability to support  $\neg S$  if  $\neg Q$  is assigned undefeated in the partial status assignment and  $\neg Q$  has no ability to support  $S$  if  $\neg Q$  is assigned defeated in the other partial status assignment. In critical-link semantics with Pollock's definition of  $\sim$ , we explained above that  $Q$ ,  $\neg Q$ , and hence  $\neg S$ , are all defeated, and then  $S$  is undefeated. With our new definition of  $\sim$  the outcome is different. For simplicity, we again assume that the strengths of reasons are at least as great as the degrees of justification of the initial node. Then the computation of  $J(\neg S, G)$  can be concluded as follows:  $J(\neg S, G) = \min\{\rho, J(\neg Q, G)\} \sim J(S, G_{\neg S}) = J(\neg Q, G) \sim J(A, G) = (J(P, G) \sim J(R, G)) \sim J(A, G)$ .

We discuss the possible degrees of justification of  $\neg Q$  and  $\neg S$ .  $\neg Q$  has ability to support  $\neg S$  iff  $J(P, G) \geq J(R, G)$ . Hence,  $J(\neg S, G) > 0$  iff  $\neg Q$  has ability to support  $\neg S$  and  $J(P, G)(1 - J(R, G)) \geq J(A, G)$ . Otherwise,  $J(\neg S, G) = 0$ . For instance, let  $J(P, G) = J(R, G) = 0.8$ ,  $J(A, G) = 0.1$  and the reason-strengths are equally strong:  $\rho = 0.9$ , then  $J(\neg Q, G) = \min\{\rho, J(P, G)\} \sim J(Q, G_{\neg Q}) = J(P, G) \sim J(R, G) = 0.16$ ;  $J(Q, G) = \min\{\rho, J(R, G)\} \sim J(\neg Q, G_Q) = J(R, G) \sim J(P, G) = 0.16$ ;  $J(\neg S, G) = \min\{\rho, J(\neg Q, G)\} \sim J(S, G_{\neg S}) = J(\neg Q, G) \sim J(A, G) = 0.144$ .

Apparently,  $\neg Q$  has the power to support  $\neg S$  and  $\neg S$  therefore has the ability to defeat or support another nodes. Moreover, if we let  $J(P, G) < J(R, G)$  or  $J(P, G)(1 - J(R, G)) < J(A, G)$ , the degree of justification of  $\neg S$  equals 0.

#### 4. The new semantics applied to the $ASPIC^+$ framework

The idea of critical-link semantics with variable degrees of justification is a general theory and can be applied in other argumentation formalisms as well. In this section, we discuss how the new semantics can be combined with the  $ASPIC^+$  argumentation framework. This framework [9,10,16–18] has been shown to capture a number of other approaches to structured argumentation, such as assumption-based argumentation [3], forms of classical argumentation [6] and Carneades [5]. Therefore, a discussion of how the above ideas can be incorporated in  $ASPIC^+$  arguably increases the generality of the present contributions. It should be noted, however, that in the present section we modify  $ASPIC^+$ 's semantics, by replacing its Dung-based semantics with a semantics that allows for variable degrees of justification.

The  $ASPIC^+$  framework is like Pollock's approach based on the general idea that arguments are constructed by chaining reasons into trees and it also assumes both deductive and defeasible reasons (unlike in Pollock's work,  $ASPIC^+$  also allows for premise attack). However,  $ASPIC^+$  does not define the notion of an inference graph but defines a set of arguments plus binary relations of attack and defeat, where defeat is a subset of attack by applying a preference relation between arguments to resolve attacks. Attack and defeat are thus not defined between argument lines but between arguments in their entirety. In essence, an argument  $A$  attacks/defeats an argument  $B$  if it directly attacks/defeats the last step of some subargument  $B'$  of  $B$  (where direct attack and defeat are defined as above for argument lines). The main technical problem then is to modify the above definitions by taking only the direct attackers/defeaters of an argument into account when computing the argument's degree of justification.

We now sketch how this can be done, where we assume that the degree of justification of one argument equals the degree of justification of its conclusion. Firstly, we introduce the  $ASPIC^+$  framework with variable degrees of justification.

#### 4.1. $ASPIC^+$ framework with variable degrees

The  $ASPIC^+$  framework assumes an unspecified logical language  $\mathcal{L}$  with a binary contrariness relation and defines arguments as inference trees formed by applying strict or defeasible inference rules of the form  $\varphi_1, \dots, \varphi_n \rightarrow \varphi$  and  $\varphi_1, \dots, \varphi_n \Rightarrow \varphi$ , where  $\varphi_1, \dots, \varphi_n$  are the *antecedents* and  $\varphi$  the *consequent*  $\varphi$  of the rule. The framework applies to any set of strict and defeasible inference rules formulated over  $\mathcal{L}$ . The inference rules can be used to express domain-specific knowledge but also to capture general patterns of reasoning.

Informally, that an inference rule is strict means that if its antecedents are accepted, then its consequent must be accepted *no matter what*, while that an inference rule is defeasible means that if its antecedents are accepted, then its consequent must be accepted *if there are no good reasons not to accept it*. In other words, if an inference rule is strict, then it is rationally impossible to accept its antecedents while refusing to accept its consequent, while if an inference rule is defeasible, it is rationally possible to accept its antecedents but not its consequent.

We now present a modified definition of  $ASPIC^+$ 's notion of an argumentation system,<sup>2</sup> allowing for degrees of strengths of inference rules. We assign real number 1 to the strength of strict inference rules and assume that the defeasible inference rules have a strength with which is any real number belonging to  $[0, 1)$ .

**Definition 19** (Argumentation system). An *argumentation system*  $AS$  is a tuple  $AS = (\mathcal{L}, -, \mathcal{R}, \nu, n)$  where

- $\mathcal{L}$  is a logical language.
- $-$  is a function from  $\mathcal{L}$  to  $2^{\mathcal{L}}$ , such that:
  - \*  $\varphi$  is a contrary of  $\psi$ , if  $\varphi \in \overline{\psi}$ ,  $\psi \notin \overline{\varphi}$ ;
  - \*  $\varphi$  is a contradictory of  $\psi$  (denoted by ' $\varphi = -\psi$ '), if  $\varphi \in \overline{\psi}$ ,  $\psi \in \overline{\varphi}$ ;
  - \* each  $\varphi \in \mathcal{L}$  has at least one contradictory.
- $\mathcal{R} = \mathcal{R}_s \cup \mathcal{R}_d$  is a set of strict ( $\mathcal{R}_s$ ) and defeasible ( $\mathcal{R}_d$ ) inference rules, such that  $\mathcal{R}_s \cap \mathcal{R}_d = \emptyset$ ;
- $\nu$  is a function that assigns the degree of support from antecedent to consequent in a strict or defeasible inference, modeled as:  $\nu(\delta) : \delta \rightarrow [0, 1]$ , where  $\delta \in \mathcal{R}$  and  $\nu(\delta_0) = 1$ , if  $\delta_0 \in \mathcal{R}_s$ ;
- $n : \mathcal{R}_d \rightarrow \mathcal{L}$  is a naming convention for defeasible rules.

Arguments are constructed from a knowledge base  $KB$ , which consists of well-formed formulas to which a strength is assigned.

**Definition 20** (Knowledge base). A *knowledge base*  $KB$  in an argumentation system  $(\mathcal{L}, -, \mathcal{R}, \nu, n)$  is a pair  $(\mathcal{K}, \eta)$  where

- $\mathcal{K} \subseteq \mathcal{L}$  consisting of two disjoint subsets  $\mathcal{K}_n$  (the axioms) and  $\mathcal{K}_p$  (the ordinary premises);
- $\eta$  is a function that assigns degrees of acceptability to elements of the knowledge base, modeled as  $\eta(A) : A \rightarrow [0, 1]$ , where  $A \in \mathcal{K}$  and  $\eta(A_0) = 1$ , if  $A_0 \in \mathcal{K}_n$ .

Arguments can be constructed step-by-step by chaining inference rules into directed acyclic graphs. Arguments thus contain subarguments, which are the structures that support intermediate conclusions

<sup>2</sup>In fact, we consider the special case where  $ASPIC^+$ 's contrariness function over  $\mathcal{L}$  corresponds to negation.

(plus the argument itself and its premises as limiting cases). In what follows, for a given argument  $\text{Prem}$  returns  $A$ 's premises,  $\text{Conc}$  returns  $A$ 's conclusion,  $\text{Sub}$  returns all of  $A$ 's sub-arguments,  $\text{Rules}$  returns all rules in  $A$ , and  $\text{TopRule}(A)$  returns the last rule applied in  $A$ .

**Definition 21** (Argument). An argument  $A$  on the basis of a knowledge base  $(\mathcal{K}, \eta)$  in an argumentation  $(\mathcal{L}, -, \mathcal{R}, v, n)$  is:

- $\varphi$  if  $\varphi \in \mathcal{K}$  with:  $\text{Prem}(A) = \{\varphi\}$ ;  $\text{Conc}(A) = \varphi$ ;  $\text{Sub}(A) = \{\varphi\}$ ;  $\text{Rules}(A) = \emptyset$ .
- $A_1, \dots, A_n \rightarrow / \Rightarrow \psi$  if  $A_1, \dots, A_n$  are arguments such that there exists a strict/defeasible rule  $\text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow / \Rightarrow \psi$  in  $\mathcal{R}_d/\mathcal{R}_s$ .

$$\begin{aligned} \text{Prem}(A) &= \text{Prem}(A_1) \cup \dots \cup \text{Prem}(A_n); \\ \text{Conc}(A) &= \psi; \\ \text{Sub}(A) &= \text{Sub}(A_1) \cup \dots \cup \text{Sub}(A_n) \cup \{A\}; \\ \text{Rules}(A) &= \text{Rules}(A_1) \cup \dots \cup \text{Rules}(A_n) \cup \{\text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow / \Rightarrow \psi\}; \\ \text{DefRules}(A) &= \{r \mid r \in \text{Rules}(A), r \in \mathcal{R}_d\}; \\ \text{Toprule}(A) &= \{\text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow / \Rightarrow \psi\}. \end{aligned}$$

An argument is *strict* if all its inference rules are strict and *defeasible* otherwise, and it is *firm* if all its premises are in  $\mathcal{K}_n$  and *plausible* otherwise. Next, we define the computation strength of arguments in accordance with Pollock's analysis on argument-strengths in [12].

**Definition 22** (Argument strength).  $\mathcal{V}$  is a function to evaluate the strength of an argument with conditions as follows:

- if  $A \in \mathcal{K}$ , then  $\mathcal{V}(A) = \eta(A)$ ;
- if  $A$  is the form  $A_1, \dots, A_n \rightarrow / \Rightarrow \psi$ , then  $\mathcal{V}(A) = \min\{\mathcal{V}(A_1), \dots, \mathcal{V}(A_n), v(n(\text{Conc}(A_1), \dots, \text{Conc}(A_n) \rightarrow / \Rightarrow \psi))\}$ .

**Definition 23** (Maximal proper subargument). Argument  $A$  is a *maximal proper subargument* of  $B$  iff  $A$  is a subargument of  $B$  and there does not exist any proper subargument  $C$  of  $B$  such that  $A$  is a proper subargument of  $C$ .

**Definition 24** (Argumentation theories). An *argumentation theory* is a triple  $AT = (AS, KB, \mathcal{V})$  where  $AS$  is an argumentation system,  $KB$  is a knowledge base in  $AS$  and  $\mathcal{V}$  is the strength of every argument on the basis of  $KB$  in  $AS$ .

Arguments can be attacked in three ways: attacking a conclusion of a defeasible inference, attacking the defeasible inference itself, or attacking a premise.

**Definition 25** ( $ASPIC^+$  attacks).  $A$  attacks  $B$  iff  $A$  undercuts, rebuts or undermines  $B$ , where:

- $A$  undercuts argument  $B$  (on  $B'$ ) iff  $\text{Conc}(A) \in -n(r)$  for some  $B' \in \text{Sub}(B)$  with defeasible top rule  $r$ .
- $A$  rebuts argument  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = -\varphi$  for some  $B' \in \text{Sub}(B)$  of the form  $B_1'', \dots, B_n'' \Rightarrow \varphi$ .
- Argument  $A$  undermines  $B$  (on  $B'$ ) iff  $\text{Conc}(A) = -\varphi$  for some  $B' = \varphi, \varphi \in \text{Prem}(B) \cap \mathcal{K}_p$ .



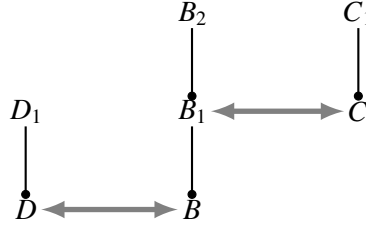


Fig. 6. Presumptive defeat.

We next give some new definitions that are useful in our modification of  $ASPIC^+$ .

**Definition 26** (Direct attacks). Argument  $A$  *directly attacks* argument  $B$  iff  $A$  rebuts or undercuts or undermines  $B$  on  $B$ ; Otherwise  $A$  indirectly attacks  $B$ .

**Corollary 1.** Argument  $A$  *directly attacks* argument  $B$  iff

- (1)  $\text{Conc}(A) = \neg \text{Conc}(B)$ , where  $B$  is not an atomic argument; or
- (2)  $\text{Conc}(A) = \neg \varphi$ , where  $\varphi \in \text{Prem}(B) \cap \mathcal{K}_p$  and  $B$  is an atomic argument; or
- (3)  $\text{Conc}(A) = \neg r$  for  $\text{TopRule}(B) = r$  and  $r$  is defeasible.

We further say that (1)  $A$  directly rebuts  $B$  or  $A$  is the direct rebutter of  $B$  iff they satisfy the first condition, (2)  $A$  directly undermines  $B$  or  $A$  is the direct underminer of  $B$  iff they satisfy the second condition, and (3)  $A$  directly undercuts  $B$  or  $A$  is the direct undercutter of  $B$  iff they satisfy the third condition.

In all previous publications on the  $ASPIC^+$  framework the definition of attack was combined with a preference ordering on arguments to yield a notion of defeat. Then the semantics was defined by regarding the set of all arguments that can be constructed plus the defeat relation as an abstract argumentation framework in the sense of [2].

Thus any semantics can be applied to arguments in an  $ASPIC^+$  framework and it can also be used to define the degrees of justification of arguments and their conclusions after it's extended with gradual degrees. We will not link the semantics of  $ASPIC^+$ 's argumentation theories with gradual values to Dung's theory, but instead of a gradual graph with two kinds of links.

#### 4.2. Modifying the $ASPIC^+$ framework with variable degrees of justification

Next we will discuss the computation of degrees of justification in  $ASPIC^+$ , using the new notion of an argument graph. We regard the degree of justification of an argument as the variable degree for accepting or rejecting the argument from a cognitive perspective. We further assume that the degree of justification of one argument equals the degree of justification of its conclusion.

**Definition 27.** An *argument graph*  $G$  is a labeled, finite, directed graph, consisting of argument nodes and attacking links indicating attacking relationships between argument nodes and proper subargument links indicating connecting subargument relationships between an argument and its proper superarguments.

The attacking links relate their roots to their targets and the root of an attacking link is an attacker in the graph, while the proper subargument links relate their roots to their targets and the root is the proper

subargument of its target or the target is the proper superargument of its root in graph. In the diagrams of argument graphs, arguments are displayed as dots, attacking links are indicated using ordinary arrowheads, while proper subargument links are indicated using closed-dot arrowheads. The initial arguments in  $G$  can be defined as follows:

**Definition 28.** An argument is *initial* in  $G$  iff it is not the target of any attacking link or proper subargument link.

**Example 1.** Consider and Pollock's inference graph in Fig. 3. We assume arguments in  $ASPIC^+$  framework as  $B = [B_1 \Rightarrow \neg S]$ ;  $B_1 = [B_2 \Rightarrow \neg Q]$ ;  $B_2 = [P]$ ;  $C = [C_1 \Rightarrow Q]$ ;  $C_1 = [R]$ ;  $D = [D_1 \Rightarrow S]$ ;  $D_1 = [A]$ . We show the arguments in Fig. 6. Note that  $C$  directly rebuts  $B_1$  and indirectly rebuts  $B$ ,  $B$  directly rebuts  $D$ . Moreover, nodes  $B_2$ ,  $C_1$  and  $D_1$  are initial arguments.

**Definition 29.** An argument path  $P(A, B)$  from argument  $A$  to argument  $B$  in graph  $G$  is a sequence of attacking links and proper subargument links  $\langle L_1, \dots, L_n \rangle$ , such that there exist arguments  $B_1, \dots, B_{n-1}$  with  $L_1 = (A, B_1)$ ,  $L_{i+1} = (B_i, B_{i+1})$ , and  $L_n = (B_{n-1}, B)$ , where  $(A, B)$  means the attacking link or proper subargument link from node  $A$  to node  $B$ .

For instance, in Fig. 6,  $P(D_1, B) = \langle (D_1, D), (D, B_1) \rangle$  and  $P(C_1, B) = \langle (C_1, C), (C, B_1), (B_1, B) \rangle$ .

**Definition 30.** An argument-path is a *circular path* from an argument  $A$  to itself iff there exists an argument path  $P(A, A)$ . Moreover, (1) An argument path is an *even cycle* iff the argument path is a circular path and the length  $n = 2k$ . (2) An argument path is an *odd cycle* iff the argument path is a circular path and the length  $n = 2k + 1$ .

For instance, in Fig. 6,  $P(D, D) = \langle (D, B), (B, D) \rangle$  and  $P(C, C) = \langle (C, B_1), (B_1, C) \rangle$  are even cycles.

**Definition 31.** Argument  $A$  is in a *circular path from an argument  $B$  to  $B$  itself* iff there exists at least one circular path from argument  $B$  to  $B$  itself containing argument  $A$ .

For instance, let  $P(A, A) = \langle (A, B), (B, C), (C, D), (D, A) \rangle$ , where  $(A, B)$  and  $(C, D)$  are attack links,  $(B, C)$  and  $(D, A)$  are proper subargument links. Then  $C$  is in the circular path  $P(A, A)$  from  $A$  to  $A$ .

Next we will make our approach simpler than Pollock's by defining the notions of a basic set and its extension instead of the notions of node-dependent and node-critical links.

**Definition 32.** The notions of basic set and critical extension can be defined as follows:

1. A set of attack links is a *basic set* of argument  $A$  in graph  $G$  iff removing all members of the set suffices to cut all cycles from  $A$  to  $A$ .
2. A set of attack links is a *critical extension* of argument  $A$  in graph  $G$  iff it is a minimum basic set of argument of  $A$  in graph  $G$ .

**Proposition 1.** For any argument  $A$  in a circular path, there exists at least one basic set of  $A$ .

**Proof.** Suppose for contradiction that there is no basic set of argument  $A$ . If there is no basic set of argument  $A$ , then there does not exist any link set such that removing all members suffices to cut the circular paths from  $A$  to  $A$ . But clearly removing all members of the set containing all links in the path can cut the circular paths. Contradiction.  $\square$

**Corollary 2.** *For any attack link  $L$  in a circular path  $P$ , there exists at least one critical extension containing  $L$ .*

**Proof.** Suppose for contradiction that there exists at least one critical extension containing  $L$ . If there is no critical extension containing  $L$ , then there is no basic set, contradicting Proposition 1.  $\square$

**Corollary 3.** *If an attack link does not occur on any circular path, then it does not belong to any critical extension.*

**Proof.** It is equivalent to prove that if an attack link belongs to a critical extension, then it occurs on a circular path. Assume attack link  $L$  belongs to a critical extension and it does not occur on any circular path, by definition 32, then there does not exist basic set, hence there is no critical extension contains  $L$ . Contradiction.  $\square$

**Definition 33.** Given a graph  $G$ , the new graph  $G_A$  is the argument-graph that results from removing all members in all critical extensions in graph  $G$  and making all arguments  $B_1, \dots, B_n$  which are not in a defeat cycle initial with  $J(B_i, G_A) = J(B_i, G)$ .

Note that Pollock only has two kinds of defeaters, namely rebuttals (attacking a conclusion) and undercutters (attacking an inference rule), so there is no computation for the defeater on premises. However, in  $ASPIC^+$ , a third way of argument attacking, namely premise attack or “undermining” has been added. So we add a calculation for the arguments which are attacked by underminers. The new computation for the degrees of justification of arguments in a new argument graph can be defined as follows:

**Definition 34.** The computation of Justification ( $JC$ ) can be defined as follows:

- (1) If argument  $A$  is initial in  $G$ ,  $J(A, G) = \mathcal{V}(A)$ .
- (2) If argument  $A$  is initial in  $G_A$ , and  $B_1, \dots, B_n$  are the direct underminers of  $A$ , then  $J(A, G) = J(A, G_A) \sim \max\{J(B_1, G_A), \dots, J(B_n, G_A)\}$ .
- (3) If argument  $A$  is not initial in  $G$  and  $G_A$ , and  $A_1, \dots, A_n$  are the maximal proper subarguments of  $A$ , and  $v(r) = \rho$  is the strength of  $Toprule(A) = r$ .  $B_1, \dots, B_i$  are the direct undercutters of  $A$  and  $B_{i+1}, \dots, B_m$  are the direct rebutters of  $A$  in a circular path  $P(A, A)$ , then  $J(A, G) = \min\{(\rho \sim \max\{J(B_1, G), \dots, J(B_i, G)\}), J(A_1, G), \dots, J(A_n, G)\} \sim \max\{J(B_{i+1}, G_A), \dots, J(B_m, G_A)\}$ .

We define  $x \sim y = x(1 - y)$ , if  $0 \leq y \leq x \leq 1$ , otherwise,  $x \sim y = 0$  and  $\max\{\emptyset\} = 0$ . The computation  $JC$  is for arguments attacked both by direct undercutters, direct rebuttals and direct underminers in cycles.  $JC(3)$  unites and double counts the computation for arguments only attacked by direct undercutters which are not in cycles and the computation for arguments only attacked by direct rebuttals.

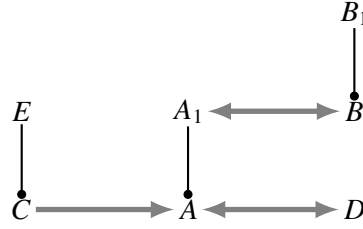


Fig. 7. Collaborative defeat.

**Example 2.** We assume the arguments in Fig. 7 as  $A_1 = [q]$ ,  $B_1 = [p]$ ,  $B = [B_1 \Rightarrow \neg q]$  with  $\text{DefRules}(B) = \{d_1 : p \Rightarrow \neg q\}$ ,  $A = [A_1 \Rightarrow s]$  with  $\text{DefRules}(A) = \{d_2 : q \Rightarrow s\}$ ,  $C = [E \Rightarrow \neg n(q \Rightarrow s)]$  with  $\text{DefRules}(C) = \{d_3 : r \Rightarrow \neg n(q \Rightarrow s)\}$ ,  $D = [\neg s]$ ,  $E = [r]$ . We show the degrees of justification of these arguments as follows:

- $B_1$  and  $E$  are initial arguments in  $G$ , then from  $JC(1)$ , we have  $J(B_1, G) = \mathcal{V}(B_1) = \eta(p)$  and  $J(E, G) = \mathcal{V}(E) = \eta(r)$ .
- $B$  directly undermines  $A_1$  and  $A_1$  is initial in  $G_A$ ,  $A$  directly undermines  $D$  and  $D$  is initial in  $G_D$ , then from  $JC(2)$ , we have  $J(A_1, G) = J(A_1, G_{A_1}) \sim J(B, G_{A_1})$  and  $J(D, G) = J(D, G_D) \sim J(A, G_D)$ .
- $A_1$  directly rebuts  $B$  and  $B$  is not initial in  $G$  and  $G_B$ , then from  $JC(3)$ , we have  $J(B, G) = \min\{\rho(d_1), J(B_1, G)\} \sim J(A_1, G_B)$ .
- $D$  directly rebuts  $A$  and  $C$  directly undercuts  $A$ , then from  $JC(3)$ , we have  $J(A, G) = \min\{(\rho(d_2) \sim J(C, G)), J(A_1, G)\} \sim J(D, G_A)$ .
- $C$  is free from attacking, so  $J(C, G) = \min\{\rho(d_3), J(E, G)\}$ .

**Example 3** (Example 1 continued). ( $JC$ ) can also be illustrated with real values. We consider the arguments in Example 1. Let  $\eta(P) = \eta(R) = 0.8$ ,  $\eta(A) = 0.1$  and the reasons are equally strong:  $\rho = 0.9$ . From  $JC(1)$ , it follows  $J(B_2, G) = J(C_1, G) = 0.8$  and  $J(D_1, G) = 0.1$ . It's easy to have that  $C$  directly rebuts  $B_1$ , then from  $JC(3)$ , it follows  $J(B_1, G) = \min\{\rho, J(B_2, G)\} \sim J(C, G_{B_1}) = J(B_2, G) \sim J(C_1, G) = 0.16$ ; We also have  $B$  directly rebuts  $D$ , then from  $JC(3)$ , it follows  $J(B, G) = \min\{\rho, J(B_1, G)\} \sim J(C, G_B) = J(B_1, G) \sim J(D_1, G) = 0.144$ ; Similar,  $J(B, G_D) = 0.16$ ;  $J(D, G) = \min\{\rho, J(D_1, G)\} \sim J(B, G_D) = J(D_1, G) \sim J(B, G_D) = 0$ .

## 5. Conclusion

In this paper we studied the modelling of variable degrees of justification in argumentation. We pointed out some arguably counter-intuitive consequences of Pollock's critical-link semantics with variable degrees of justification and then presented some modifications that avoid these outcomes. Moreover, to illustrate the generality of Pollock's approach and our modifications, we also discussed how they can be combined with the  $ASPIC^+$  framework. In future work we aim to investigate the properties of our definitions and to study their application to realistic examples, including problems of legal reasoning with evidence.

While our contributions in this paper still need to be further developed in these and other ways, we hope we have contributed to the study of an important but often neglected aspect of natural argumentation. In most current approaches, the justification status of arguments and conclusions is an all-or-nothing

affair, but in many realistic applications, such as legal reasoning about evidence or other applications of epistemic reasoning, it is natural to regard arguments or conclusions as justified to variable degrees.

## Acknowledgements

We thank the anonymous reviewers for their useful comments on the earlier versions of this paper. Bin Wei was supported by the Humanity and Social Science Youth foundation of Ministry of Education of China (15YJCZH182), the Chinese MOE Project of Key Research Institute of Humanities and Social Sciences at Universities (15JJD720014) and by the Scientific and Technological Research Program of Chongqing Municipal Education Commission (KJ1500103).

## References

- [1] P. Baroni and M. Giacomin, Solving semantic problems with odd-length cycles in argumentation, in: *Proceedings of the 7th European Conference on Symbolic and Quantitative Approaches to Reasoning with Uncertainty*, Springer-Verlag, 2003, pp. 440–451. doi:[10.1007/978-3-540-45062-7\\_36](https://doi.org/10.1007/978-3-540-45062-7_36).
- [2] P.M. Dung, On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and  $n$ -person games, *Artificial Intelligence* **77** (1995), 321–377. doi:[10.1016/0004-3702\(94\)00041-x](https://doi.org/10.1016/0004-3702(94)00041-x).
- [3] P.M. Dung, P. Mancarella and F. Toni, Computing ideal sceptical argumentation, *Artificial Intelligence* **171** (2007), 642–674. doi:[10.1016/j.artint.2007.05.003](https://doi.org/10.1016/j.artint.2007.05.003).
- [4] A. Garcia and G. Simari, Defeasible logic programming: An argumentative approach, *Theory and Practice of Logic Programming* **4** (2004), 95–138. doi:[10.1017/s1471068403001674](https://doi.org/10.1017/s1471068403001674).
- [5] T.F. Gordon, H. Prakken and D.N. Walton, The Carneades model of argument and burden of proof, *Artificial Intelligence* **171** (2007), 875–896. doi:[10.1016/j.artint.2007.04.010](https://doi.org/10.1016/j.artint.2007.04.010).
- [6] N. Gorogiannis and A. Hunter, Instantiating abstract argumentation with classical-logic arguments: Postulates and properties, *Artificial Intelligence* **175** (2011), 1479–1497. doi:[10.1016/j.artint.2010.12.003](https://doi.org/10.1016/j.artint.2010.12.003).
- [7] H. Jakobovits and D. Vermeir, Robust semantics for argumentation frameworks, *Journal of Logic and Computation* **9** (1999), 215–261. doi:[10.1093/logcom/9.2.215](https://doi.org/10.1093/logcom/9.2.215).
- [8] D. Makinson and K. Schlechta, Floating conclusions and zombie paths, *Artificial Intelligence* **48** (1991), 199–209. doi:[10.1016/0004-3702\(91\)90061-n](https://doi.org/10.1016/0004-3702(91)90061-n).
- [9] S. Modgil and H. Prakken, A general account of argumentation with preferences, *Artificial Intelligence* **195** (2013), 361–397. doi:[10.1016/j.artint.2012.10.008](https://doi.org/10.1016/j.artint.2012.10.008).
- [10] S.J. Modgil and H. Prakken, Revisiting preferences and argumentation, in: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, 2011, pp. 1021–1026.
- [11] J.L. Pollock, Defeasible reasoning, *Cognitive Science* **11** (1987), 481–518. doi:[10.1207/s15516709cog1104\\_4](https://doi.org/10.1207/s15516709cog1104_4).
- [12] J.L. Pollock, Defeasible reasoning with variable degrees of justification, *Artificial Intelligence* **133** (2002), 233–282. doi:[10.1016/s0004-3702\(01\)00145-x](https://doi.org/10.1016/s0004-3702(01)00145-x).
- [13] J.L. Pollock, A recursive semantics for defeasible reasoning, in: *Argumentation in Artificial Intelligence*, I. Rahwan and G. Simari, eds, Springer, Berlin, 2009, pp. 173–197. doi:[10.1007/978-0-387-98197-0\\_9](https://doi.org/10.1007/978-0-387-98197-0_9).
- [14] J.L. Pollock, Defeasible reasoning and degrees of justification, *Argument and Computation* **1**(1) (2010), 7–22. doi:[10.1080/19462161003728921](https://doi.org/10.1080/19462161003728921).
- [15] J.L. Pollock, *Cognitive Carpentry: A Blueprint for How to Build a Person*, MIT Press, Cambridge, MA, 1995.
- [16] H. Prakken, An abstract framework for argumentation with structured arguments, *Argument and Computation* **1**(2) (2010), 93–124. doi:[10.1080/19462160903564592](https://doi.org/10.1080/19462160903564592).
- [17] B. van Gijzel and H. Prakken, Relating Carneades with abstract argumentation via the ASPIC+ framework for structured argumentation, *Argument and Computation* **3**(1) (2012), 21–47. doi:[10.1080/19462166.2012.661766](https://doi.org/10.1080/19462166.2012.661766).
- [18] B. van Gijzel and H. Prakken, Relating Carneades with abstract argumentation, in: *Proceedings of the 22nd International Joint Conference on Artificial Intelligence (IJCAI 2011)*, 2011, pp. 1113–1119.